WHAT IS CLAIMED IS:

1.　　A method for allelic classification, the method comprising:

acquiring intensity information for a plurality of samples wherein the intensity information comprises a first intensity component associated with a first allele and a second intensity component associated with a second allele;

evaluating the intensity information for each of the plurality of samples to identify one or more data clusters, each cluster associated with a discrete allelic combination and determined, in part, by comparing the first intensity component relative to the second intensity component;

generating a likelihood model that predicts the probability that a selected sample will reside within a particular data cluster based upon its intensity information; and

applying the likelihood model to each of the plurality of samples to determine its associated allelic composition.

2.　　The method of Claim 1, wherein the likelihood model comprises a model-fit probability assessment that estimates confidence in the likelihood model itself and assesses how well a selected sample and its respective intensity information fit the model.

3.　　The method of Claim 1, wherein the likelihood model comprises an in-class probability assessment that estimates the probability that a selected cluster identifies a selected sample and its respective intensity information.

4.　　The method of Claim 1, wherein the likelihood model comprises an a posteriori probability assessment that estimates the probability of a selected sample and its respective intensity information belonging to an assigned cluster.

5.	The method of Claim 1, wherein the data clusters comprise at least three discrete clusters each associated with a different allelic classification.

6.	The method of Claim 5, wherein the data clusters comprise a first cluster type associated with a first homozygous allelic classification.

7.	The method of Claim 6, wherein the data clusters comprise a second cluster type associated with a first heterozygous allelic classification.

8.	The method of Claim 7, wherein the data clusters comprise a third cluster type associated with a second homozygous allelic classification.

9.	The method of Claim 1, wherein the allelic classification is used to perform a mutational analysis of one or more samples.

10.	The method of Claim 1, wherein the allelic classification is used to perform a single nucleotide polymorphism analysis of one or more samples.

11.	The method of Claim 1, wherein the geneotype for one or more samples is identified by performing the allelic classification.

12.	The method of Claim 1, wherein the intensity information for the plurality of clusters is normalized.

13.	The method of Claim 1, wherein the plurality of samples comprise at least one "no template control" sample and associated intensity information that is used for the purposes of sample scaling.

14.	The method of Claim 1, wherein the likelihood model is generated in an iterative manner to refine the likelihood model.

15.   The method of Claim 14, wherein two or more iterations are used to generate a refined likelihood model.

16.   The method of Claim 14, wherein refinement of the likelihood model is performed by identifying outlier samples and removing these samples prior to further likelihood model generation to generate a refined likelihood model.

17.   The method of Claim 14, wherein refinement of the likelihood model comprises performing a data resampling operation wherein a subset of the plurality of samples are used to generate the refined likelihood model.

18.   The method of Claim 1, wherein the first and second intensity components of the intensity information comprise fluorescence intensities associated with discrete markers or labels.

19.   The method of Claim 1, wherein the intensity information for each sample is acquired from a dual-label amplification protocol.

20.   The method of Claim 19, wherein the dual-label amplification protocol comprises a Taqman or SNPlex protocol.

21.   The method of Claim 1, wherein the intensity information for each sample is acquired from an array-based detection protocol.

22.   A method for clustering analysis, the method comprising:
       identifying a sample set comprising a plurality of data points, each data point having an angular value representative of an association between a first and a second intensity component;

generating a likelihood model and associated parameter set wherein the angular values of the data points are used in determining the appropriate parameters to be used in the likelihood model and wherein the efficacy of the likelihood model is assessed by evaluating the probability the likelihood model properly identifies selected data points in the sample set;

applying the likelihood model to the plurality of data points within the sample set and grouping the data points into discrete clusters; and

associating a selected classification with each discrete cluster and its component data points.

23.   The method of Claim 22, wherein the clustering analysis is used in allelic classification.

24.   The method of Claim 23, wherein the allelic classification comprises identifying the discrete clusters representing a homozygous allelic classification or a heterozygous allelic classification and associating the data points of a particular cluster with the identified allelic classification.

25.   The method of Claim 23, wherein at least three discrete clusters exist which correspond to a first homozygous allelic classification, a second homozygous allelic classification, and a first heterozygous allelic classification.

26.   The method of Claim 22, wherein the clustering analysis is used to perform mutational analysis.

27.   The method of Claim 22, wherein the clustering analysis is used to perform single nucleotide polymorphism analysis.

28.   The method of Claim 22, wherein the likelihood model and associated parameter set are evaluated using a probability assessment the estimates

confidence in the likelihood model itself and assesses how well as selected data point fits the model using the associated parameter set.

29.   The method of Claim 22, wherein the likelihood model and associated parameter set are evaluated using a probability assessment that estimates the probability that a selected cluster properly identifies a selected data point associated with the cluster.

30.   The method of Claim 22, wherein the likelihood model and associated parameter set are evaluated using a probability assessment that estimates the probability that a selected data point belongs to the cluster to which it is grouped.

31.   The method of Claim 22, wherein the likelihood model and associated parameter set are generated in an iterative manner wherein one or more data points are excluded from the model and parameter analysis and a second refined model and parameter set is generated using the remaining data points.

32.   The method of Claim 31, wherein the excluded data points comprise outlier data points which reside beyond a defined cluster threshold.

33.   The method of Claim 31, wherein additional refinements to the model and parameter set are performed by excluding additional data points.

34.   A method for allelic classification, the method comprising:
identifying a sample set comprising a plurality of data points each having at least two component intensity values;
evaluating the component intensity values for the plurality of data points to group the data points into one or more data clusters representative of discrete allelic classifications;

generating a likelihood function that describes the grouping of a selected data point using its component intensity value; and

associating an allelic classification with each data point using the likelihood function.

35. The method of Claim 34, further comprising performing a confidence value assessment for each data point indicative of a degree of confidence with which the allelic classification is made.

36. The method of Claim 34, further comprising a refinement operation in which at least one data point is excluded from the sample set and a refined likelihood function is generated based on the remaining data points of the sample set.

37. The method of Claim 36, wherein the at least one excluded data point comprises outlier data which resides outside of a selected grouping.

38. The method of Claim 34, wherein at least three groupings of data points are present and correspond to a first homozygous allelic classification, a second allelic classification and a first heterozygous classification.

39. The method of Claim 34, wherein the likelihood function efficacy is further evaluated based on the confidence of the likelihood model itself and how well data points fit into the model.

40. The method of Claim 34, wherein the likelihood function efficacy is further evaluated according to the probability that a selected data point belongs to the associated allelic classification.

41.   The method of Claim 34, wherein the likelihood function efficacy is further evaluated according to the probability that a selected data cluster could be associated with a particular data point.

42.   A computer readable medium having stored thereon instructions which cause a general purpose computer to perform the steps of:

acquiring experimental information for a plurality of samples wherein the experimental information comprises a first data component associated with a first allele and a second data component associated with a second allele;

evaluating the experimental information for each of the plurality of samples to identify one or more data clusters, each cluster associated with a discrete allelic combination and determined, in part, by comparing the first data component relative to the second data component;

generating a likelihood model that predicts the probability that a selected sample will reside within a particular data cluster based upon its experimental information; and

applying the likelihood model to each of the plurality of samples to determine its associated allelic composition.

43.   The computer readable medium of Claim 42, wherein the first and second data component comprise sample intensity information.

44.   The computer readable medium of Claim 43, wherein the sample intensity information is acquired following reacting each sample using a dual-label amplification protocol.

45.   The computer readable medium of Claim 44, wherein the dual-label amplification protocol comprises a Taqman or SNPlex protocol.

46.   The computer readable medium of Claim 42, wherein the likelihood model comprises a model-fit probability assessment that estimates confidence in the likelihood model itself and assesses how well a selected sample and its respective experimental information fit the model.

47.   The computer readable medium of Claim 42, wherein the likelihood model comprises an in-class probability assessment that estimates the probability that a selected cluster identifies a selected sample and its respective experimental information.

48.   The computer readable medium of Claim 42, wherein the likelihood model comprises an a posteriori probability assessment that estimates the probability of a selected sample and its respective experimental information belonging to an assigned cluster.

49.   The computer readable medium of Claim 42, wherein the data clusters comprise at least three discrete clusters each associated with a different allelic classification.

50.   The computer readable medium of Claim 49, wherein the data clusters comprise a first cluster type associated with a first homozygous allelic classification, a second cluster type associated with a first heterozygous allelic classification, and a third cluster type associated with a second homozygous allelic classification.

51.   The computer readable medium of Claim 42, wherein the data clusters comprise one or more clusters each associated with a discrete allelic classification.

52.   The computer readable medium of Claim 42, wherein the steps further comprise normalizing the experimental information.

53.   The computer readable medium of Claim 42, wherein the steps further operate in an iterative manner to refine the likelihood model.

54.   The computer readable medium of Claim 53, wherein two of more iterations are used to generate a refined likelihood model.

55.   The computer readable medium of Claim 54, wherein the likelihood model is refined by identifying outlier samples and removing these samples prior to further likelihood model generation.

56.   The computer readable medium of Claim 42, wherein the experimental information comprises angular data.

57.   The computer readable medium of Claim 56, wherein the angular data is generated by comparing the first data component with the second data component for each sample.

58.   The computer readable medium of Claim 56, wherein the angular data reflects a ratio between the first data component and the second data component for each sample.

59.   A computer readable medium having stored thereon instructions which cause a general purpose computer to perform the steps of:

    identifying a sample set comprising a plurality of data points, each data point having an angular value representative of an association between a first and a second intensity component;

    generating a likelihood model and associated parameter set wherein the angular values of the data points are used in determining the appropriate parameters to be used in the likelihood model and wherein the efficacy of the

-39-

likelihood model is assessed by evaluating the probability the likelihood model properly identifies selected data points in the sample set;

applying the likelihood model to the plurality of data points within the sample set and grouping the data points into discrete clusters; and

associating a selected classification with each discrete cluster and its component data points.

60. The computer readable medium of Claim 59, wherein the operations are used to perform allelic classification in which the discrete clusters represent a homozygous allelic classification or a heterozygous allelic classification and the data points of a particular cluster are associated with the corresponding allelic classification.

61. The computer readable medium of Claim 60, wherein at least three discrete clusters exist which correspond to a first homozygous allelic classification, a second homozygous allelic classification, and a first heterozygous allelic classification.

62. The computer readable medium of Claim 59, wherein the likelihood model and associated parameter set are evaluated using a probability assessment the estimates confidence in the likelihood model itself and assesses how well as selected data point fits the model using the associated parameter set.

63. The computer readable medium of Claim 59, wherein the likelihood model and associated parameter set are evaluated using a probability assessment that estimates the probability that a selected cluster properly identifies a selected data point associated with the cluster.

64. The computer readable medium of Claim 59, wherein the likelihood model and associated parameter set are evaluated using a probability assessment

that estimates the probability that a selected data point belongs to the cluster to which it is grouped.

65.   The computer readable medium of Claim 59, wherein the likelihood model and associated parameter set are generated in an iterative manner wherein one or more data points are excluded from the model and parameter analysis and a second refined model and parameter set is generated using the remaining data points.

66.   A computer readable medium having stored thereon instructions which cause a general purpose computer to perform the steps of:

identifying a sample set comprising a plurality of data points each having at least two component experimental values;

evaluating the component experimental values for the plurality of data points to group the data points into one or more data clusters representative of discrete allelic classifications;

generating a likelihood function that describes the grouping of a selected data point using its component experimental value; and

associating an allelic classification with each data point using the likelihood function.

67.   The computer readable medium of Claim 66, the steps further comprising performing a confidence value assessment for each data point indicative of a degree of confidence with which the allelic classification is made.

68.   The computer readable medium of Claim 66, the steps further comprising a refinement operation in which at least one data point is excluded from the sample set and a refined likelihood function is generated based on the remaining data points of the sample set.

69.   A computer-based system for performing allelic classification, the system comprising:

a database for storing experimental information for a plurality of samples, the experimental information reflecting the allelic composition of each sample;

a program which performs the operations of:

retrieving experimental information for the plurality of samples from the database wherein the experimental information comprises a first data component associated with a first allele and a second data component associated with a second allele;

evaluating the experimental information for each of the plurality of samples to identify one or more data clusters, each cluster associated with a discrete allelic combination and determined, in part, by comparing the first experimental component relative to the experimental component;

generating a likelihood model comprising a model-fit probability assessment that estimates confidence in the likelihood model itself and assesses how well a selected sample and its respective experimental information fit the model, the model further used to predict the probability that a selected sample is associated with a particular data cluster based upon its experimental information; and

applying the likelihood model to each of the plurality of samples to determine its associated allelic composition.

70.   The system of Claim 69, wherein the first and second data component comprise sample intensity information.

71.   The system of Claim 69, wherein the likelihood model comprises an in-class probability assessment that estimates the probability that a selected cluster identifies a selected sample and its respective experimental information.

72.   The system of Claim 69, wherein the likelihood model comprises an a posteriori probability assessment that estimates the probability of a selected sample and its respective experimental information belonging to an assigned cluster.

73.   The system of Claim 69, wherein the data clusters comprise at least three discrete clusters each associated with a different allelic classification.

74.   The system of Claim 73, wherein the data clusters comprise a first cluster type associated with a first homozygous allelic classification, a second cluster type associated with a first heterozygous allelic classification, and a third cluster type associated with a second homozygous allelic classification.

75.   The system of Claim 69, wherein the program further operates to normalize the experimental information.

76.   The system of Claim 69, wherein the program further operates in an iterative manner to refine the likelihood model.

77.   The system of Claim 76, wherein two of more iterations are used to generate a refined likelihood model.

78.   The system of Claim 76, wherein the program refines the likelihood model by identifying outlier samples and removing these samples prior to further likelihood model generation to generate the refined likelihood model.

79.   The system of Claim 69, wherein the experimental information comprises angular data generated by comparing the first data component with the second data component for each sample.

80.   A computer-based system for performing allelic classification, the system comprising:

a database for storing experimental information for a plurality of samples, the experimental information reflecting the allelic composition of each sample; and

a program which performs the operations of:

identifying a sample set comprising a plurality of data points, each data point having an angular value representative of an association between a first and a second intensity component;

generating a likelihood model and associated parameter set wherein the angular values of the data points are used in determining the appropriate parameters to be used in the likelihood model and wherein the efficacy of the likelihood model is assessed by evaluating the probability the likelihood model properly identifies selected data points in the sample set;

applying the likelihood model to the plurality of data points within the sample set and grouping the data points into discrete clusters; and

associating a selected classification with each discrete cluster and its component data points.

81.   The system of Claim 80, wherein the clustering analysis is used in allelic classification by identifying the discrete clusters representing a homozygous allelic classification or a heterozygous allelic classification and associating the data points of a particular cluster with the identified allelic classification.

82.   The system of Claim 81, wherein at least three discrete clusters exist which correspond to a first homozygous allelic classification, a second homozygous allelic classification, and a first heterozygous allelic classification.

83. The system of Claim 80, wherein the likelihood model and associated parameter set are evaluated using a probability assessment the estimates confidence in the likelihood model itself and assesses how well as selected data point fits the model using the associated parameter set.